What is claimed is:

1. A purified or isolated *BAP28* nucleic acid comprising at least 12 contiguous nucleotides of the nucleotide sequence of SEQ ID No 1, or the complements thereof, wherein said contiguous span comprises at least one of the following nucleotide positions of SEQ ID No 1: 1-50357, 50499-50963, 51257-52147, 52299-53234, 53394-53553, 53689-53837, 53943-54028, 54198-54740, 54896-55753, 55913-57385, 57495-58503, 58828-85946, 59355-85946, 86169-91228, and/or 91852 to 97662.

2. A purified or isolated nucleic acid encoding a BAP28 protein comprising at least 12 consecutive nucleotides of a nucleotide sequence selected from the group consisting of SEQ ID Nos 2 and 3 or the complement thereof, wherein said contiguous span comprises at least 1 of nucleotide positions 1 to 4995 of SEQ ID No 2 or 3.

3. An isolated, purified or recombinant polynucleotide consisting essentially of a contiguous span of 8 to 50 nucleotides of SEQ ID No 1 or the complement thereof, wherein a span includes a *BAP28*-related biallelic marker.

4. A purified or isolated nucleic acid according to claim 3, wherein said contiguous span comprises a *BAP28*-related biallelic marker selected from the group consisting of A1 to A58, and the complements thereof.

5. A purified or isolated nucleic acid according to claim 3, wherein said contiguous span comprises a *BAP28*-related biallelic marker selected from the group consisting of A1 to A27, A34, A37 to A41, A43 to A49, A52, and A54 to A58, and the complements thereof.

6. A polynucleotide according to claim 3, wherein said contiguous span is 18 to 35 nucleotides in length and said biallelic marker is within 4 nucleotides of the center of said polynucleotide.

7. A polynucleotide according to claim 6, wherein said polynucleotide consists of said contiguous span and said contiguous span is 25 nucleotides in length and said biallelic marker is at the center of said polynucleotide.

8. A polynucleotide according to claim 7, wherein said polynucleotide consists essentially of a sequence selected from the following sequences: P1 to P58, and the complementary sequences thereto.

9. A polynucleotide according to claim 3, wherein the 3' end of said contiguous span is located at the 3' end of said polynucleotide and said biallelic marker is present at the 3' end of said polynucleotide.

10. An isolated, purified, or recombinant polynucleotide consisting essentially of a contiguous span of 8 to 50 nucleotides of anyone of SEQ ID Nos 1, 2, or 3 or the complement thereof, wherein the 3' end of said contiguous span is located at the 3' end of said polynucleotide, and wherein the 3' end of said polynucleotide is located within 20 nucleotides upstream of a *BAP28*-related biallelic marker in said sequence.

11. A polynucleotide according to claim 10, wherein the 3' end of said polynucleotide is located 1 nucleotide upstream of a BAP28-related biallelic marker in said sequence.

12. A polynucleotide according to claim 11, wherein said polynucleotide consists essentially of a sequence selected from the following sequences: D1 to D58, and E1 to E58.

13. An isolated, purified, or recombinant polynucleotide consisting essentially of a sequence selected from the following sequences: B1 to B38 and C1 to C38.

14. An isolated, purified, or recombinant polynucleotide which encodes a polypeptide comprising a contiguous span of at least 6 amino acids of SEQ ID No 5, wherein said contiguous span includes:

- at least 1 of the amino acid positions 1 to 1629 of the SEQ ID No 5; or,
- an amino acid selected from the group consisting of an asparagine at the amino acid position 1694 of SEQ ID No 5, a valine at the amino acid position 1854 of SEQ ID No 5, an asparagine at the amino acid position 1967 of SEQ ID No 5, a glutamic acid at the amino acid position 2017 of SEQ ID No 5, and an alanine at the amino acid position 2050 of SEQ ID No 5.

15. An isolated, purified, or recombinant polynucleotide comprising a sequence selected from the group consisting of SEQ ID Nos 4, and 9-13 and the complementary sequence thereto.

16. A polynucleotide according to any one of claims 1-3, 10, 13-15 attached to a solid support.

17. An array of polynucleotides comprising at least one polynucleotide according to claim 16.

18. An array according to claim 17, wherein said array is addressable.

19. A polynucleotide according to any one of claims 1-3, 10, 13-15 further comprising a label.

20. A recombinant vector comprising a polynucleotide according to any one of claims 1-3, 10, 13-15.

21. A host cell comprising a recombinant vector according to claim 20.

22. A non-human host animal or mammal comprising a recombinant vector according to claim 20.

23. A mammalian host cell comprising a *BAP28* gene disrupted by homologous recombination with a knock out vector, comprising a polynucleotide according to any one of claims 1-3 and 14.

24. A non-human host mammal comprising a *BAP28* gene disrupted by homologous recombination with a knock out vector, comprising a polynucleotide according to any one of claims 1-3 and 14.

25. A method of genotyping comprising determining the identity of a nucleotide at a *BAP28* -related biallelic marker or the complement thereof in a biological sample.

26. A method according to claim 25, wherein said biological sample is derived from a single subject.

27. A method according to claim 26, wherein the identity of the nucleotides at said biallelic marker is determined for both copies of said biallelic marker present in said individual's genome.

28. A method according to claim 25, wherein said biological sample is derived from multiple subjects.

29. A method according to claim 25, further comprising amplifying a portion of said sequence comprising the biallelic marker prior to said determining step.

30. A method according to claim 29, wherein said amplifying is performed by PCR.

31. A method according to claim 25, wherein said determining is performed by an assay selected from the group consisting of hybridization assay, a sequencing assay, a microsequencing assay, and an enzyme-based mismatch detection assay.

32. A method of estimating the frequency of an allele of a *BAP28*-related biallelic marker in a population comprising:

a) genotyping individuals from said population for said biallelic marker according to the method of claim 25; and

b) determining the proportional representation of said biallelic marker in said population..

33. A method of detecting an association between a genotype and a trait, comprising the steps of:

a) determining the frequency of at least one *BAP28*-related biallelic marker in trait positive population according to the method of claim 32;

b) determining the frequency of at least one *BAP28*-related biallelic marker in a control population according to the method of claim 32; and

c) determining whether a statistically significant association exists between said genotype and said trait.

34. A method of estimating the frequency of a haplotype for a set of biallelic markers in a population, comprising:

a) genotyping at least one *BAP28*-related biallelic marker according to claim 27 for each individual in said population;

b) genotyping a second biallelic marker by determining the identity of the nucleotides at said second biallelic marker for both copies of said second biallelic marker present in the genome of each individual in said population; and

c) applying a haplotype determination method to the identities of the nucleotides determined in steps a) and b) to obtain an estimate of said frequency.

35. A method according to claim 34, wherein said haplotype determination method is selected from the group consisting of asymmetric PCR amplification, double PCR amplification of specific alleles, the Clark algorithm, or an expectation-maximization algorithm.

36. A method of detecting an association between a haplotype and a trait, comprising the steps of:

a) estimating the frequency of at least one haplotype in a trait positive population according to the method of claim 34;

b) estimating the frequency of said haplotype in a control population according to the method of claim 34; and

c) determining whether a statistically significant association exists between said haplotype and said trait.

37. A method according to claim 33, wherein said genotyping steps a) and b) are performed on a single pooled biological sample derived from each of said populations.

38. A method according to claim 33, wherein said genotyping steps a) and b) performed separately on biological samples derived from each individual in said populations.

39. A method according to either claim 33 or 36, wherein said trait is cancer.

40. A method according to either claim 33 or 36, wherein said control population is a trait negative population.

41. A method according to either claim 33 or 36, wherein said case control population is a random population.

42. A method of determining whether an individual is at risk of developing prostate cancer, comprising:

a) genotyping at least one *BAP28*-related biallelic marker according to the method of claim 27; and

b) correlating the result of step a) with a risk of developing prostate cancer.

43. A method according to any one of claims 25, 32, 33, 34, 36, and 42 wherein said *BAP28*-related biallelic marker is selected from the group consisting of A1 to A58 and the complements thereof.

44. A method according to claim 42, wherein said *BAP28*-related biallelic marker is selected from the following list of biallelic markers: A1, A4, 16, A30, A31, A42, A50, A51, and A53, and the complements thereof.

45. A diagnostic kit comprising a polynucleotide according to any one of claims 3-13 and 16-19.

46. An isolated, purified, or recombinant polypeptide comprising a contiguous span of at least 6 amino acids of SEQ ID No 5, wherein said contiguous span includes:

- at least 1 of the amino acid positions 1 to 1629 of the SEQ ID No 5; or,
- an amino acid selected from the group consisting of an asparagine at the amino acid position 1694 of SEQ ID No 5, a valine at the amino acid position 1854 of SEQ ID No 5, an asparagine at the amino acid position 1967 of SEQ ID No 5, a glutamic acid at the amino acid position 2017 of SEQ ID No 5, and an alanine at the amino acid position 2050 of SEQ ID No 5.

47. An isolated or purified antibody composition are capable of selectively binding to an epitope-containing fragment of a polypeptide according to claim 46, wherein said epitope comprises:

- at least 1 of the amino acid positions 1 to 1629 of the SEQ ID No 5; or,
- an amino acid selected from the group consisting of an asparagine at the amino acid position 1694 of SEQ ID No 5, a valine at the amino acid position 1854 of SEQ ID No 5, an asparagine at the amino acid position 1967 of SEQ ID No 5, a glutamic acid at the amino acid position 2017 of SEQ ID No 5, and an alanine at the amino acid position 2050 of SEQ ID No 5.

48. A method for the screening of a candidate substance interacting with a BAP28 polypeptide comprising the following steps :

a) providing a polypeptide according to claim 46;

b) obtaining a candidate substance;

c) bringing into contact said polypeptide with said candidate substance; and

d) detecting the complexes formed between said polypeptide and said candidate substance.

49. A method for screening of a candidate substance that modulated the expression of the *BAP28* gene comprising the following steps:

a) providing a recombinant cell host containing a nucleic acid, wherein said nucleic acid comprises a nucleotide sequence of the 5' regulatory region (2996 -4996 of SEQ ID No 1) or a

biologically active fragment or variant thereof located upstream a polynucleotide encoding a detectable protein:

- obtaining a candidate substance; and

- determining the ability of the candidate substance to modulate the expression levels of the polynucleotide encoding the detectable protein.

50. A computer readable medium having stored thereon a sequence selected from the group consisting of a nucleic acid code comprising one of the following:

a) a contiguous span of at least 12 nucleotides of SEQ ID No 1, wherein said contiguous span comprises at least 1of the following nucleotide positions of SEQ ID No 1: 1-50357, 50499-50963, 51257-52147, 52299-53234, 53394-53553, 53689-53837, 53943-54028, 54198-54740, 54896-55753, 55913-57385, 57495-58503, 58828-85946, 59355-85946, 86169-91228, and/or 91852 to 97662;

b) a contiguous span of at least 12 nucleotides of SEQ ID No 1 or the complement thereof, wherein said contiguous span comprises at least 1 nucleotides selected from the group consisting of the following nucleotide positions of SEQ ID No 1: 4997-5076, 5371-5544, 6121-6337, 9877-10018, 11522-11623, 12521-12661, 13453-13664, 13824-13957, 15376-15478, 16855-16965, 17378-17495, 18535-18642, 21446-21541, 21999-22087, 23036-23247, 23546-23667, 24270-24461, 26287-26470, 26611-26747, 28068-28260, 32540-32709, 33112-33270, 34586-34828, 35156-35287, 36660-36763, 36934-37077, 37803-37921, 38017-38138, 40365-40493, 42618-42848, 43452-43578, 44836-44999, 48223-48269, and 49656-49779;

c) a contiguous span of at least 12 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises at least one BAP28-related biallelic marker selected from the group consisting of A1 to A58;

d) a contiguous span of at least 12 nucleotides of a nucleic acid sequence selected from the group consisting of SEQ ID Nos 2 and 3 or the complements thereof, wherein said contiguous span comprises at least 1 of nucleotide positions 1 to 4995 of SEQ ID No 2 or 3;

e) a contiguous span of at least 12 nucleotides of a nucleic acid sequence selected from the group consisting of SEQ ID Nos 2 and 3 or the complements thereof, wherein said contiguous span comprises at least 1 of nucleotide positions 1 to 2033, 2160 to 2348 and 2676 to 4995 of SEQ ID No 2 or 3;

f) a contiguous span of at least 12 nucleotides of a nucleic acid sequence selected from the group consisting of SEQ ID Nos 1-3 or the complements thereof, wherein said contiguous span comprises at least 1 of any one of the following ranges of nucleotide positions of:

(1) SEQ ID No 1: 1-2500, 2501-5000, 5001-7500, 7501-10000, 10001-12500, 12501-15000, 15001-17500, 17501-20000, 20001-22500, 22501-25000, 25001-27500, 27501-30000, 30001-32500, 32501-35000, 35001-37500, 37501-40000, 40001-42500, 42501-45000, 45001-

47500, 47501-50000, 50001-50357, 50499-50963, 51257-52147, 52299-53234, 53394-53553, 53689-53837, 53943-54028, 54198-54740, 54896-55753, 55913-57385, 57495-58503, 58828-85946, 59355-85946, 86169-91228, and/or 91852 to 97662;

(2) SEQ ID No 2: 1 to 500, 501 to 1000, 1001 to 1500, 1501 to 2000, 2001 to 2500, 2501 to 3000, 3001 to 3500, 3501 to 4000, 4001 to 4500, 4501 to 4995, 5000 to 5500, 5501 to 6000, 6001 to 6500, and 6501 to 6782; and,

(3) SEQ ID No 3: 1 to 500, 501 to 1000, 1001 to 1500, 1501 to 2000, 2001 to 2500, 2501 to 3000, 3001 to 3500, 3501 to 4000, 4001 to 4500, 4501 to 4995, 5000 to 5500, 5501 to 6000, 6001 to 6500, 6501 to 7000, 7001 to 7500, 7501 to 7932; and

g) a nucleotide sequence selected from the group consisting of SEQ ID Nos 4, and 9-13; and,

h) a nucleotide sequence complementary to any one of the preceding nucleotide sequences.

51. A computer readable medium having stored thereon a sequence consisting of a polypeptide code comprising a contiguous span of at least 6 amino acids of SEQ ID No 5, wherein said contiguous span includes:

- at least 1 of the amino acid positions 1 to 1629 of the SEQ ID No 5; or,
- an amino acid selected from the group consisting of an asparagine at the amino acid position 1694 of SEQ ID No 5, a valine at the amino acid position 1854 of SEQ ID No 5, an asparagine at the amino acid position 1967 of SEQ ID No 5, a glutamic acid at the amino acid position 2017 of SEQ ID No 5, and an alanine at the amino acid position 2050 of SEQ ID No 5.

52. A computer system comprising a processor and a data storage device wherein said data storage device a computer readable medium according to with claim 50 or 51.

53. A computer system according to claim 52, further comprising a sequence comparer and a data storage device having reference sequences stored thereon.

54. A computer system of Claim 53 wherein said sequence comparer comprises a computer program which indicates polymorphisms.

55. A computer system of Claim 52 further comprising an identifier which identifies features in said sequence.

56. A method for comparing a first sequence to a reference sequence, comprising the steps of:

reading said first sequence and said reference sequence through use of a computer program which compares sequences; and

determining differences between said first sequence and said reference sequence with said computer program,

wherein said first sequence is selected from the group consisting of a nucleic acid code comprising one of the following:

a) a contiguous span of at least 12 nucleotides of SEQ ID No 1, wherein said contiguous span comprises at least 1of the following nucleotide positions of SEQ ID No 1: 1-50357, 50499-50963, 51257-52147, 52299-53234, 53394-53553, 53689-53837, 53943-54028, 54198-54740, 54896-55753, 55913-57385, 57495-58503, 58828-85946, 59355-85946, 86169-91228, and/or 91852 to 97662;

b) a contiguous span of at least 12 nucleotides of SEQ ID No 1 or the complement thereof, wherein said contiguous span comprises at least 1 nucleotides selected from the group consisting of the following nucleotide positions of SEQ ID No 1: 4997-5076, 5371-5544, 6121-6337, 9877-10018, 11522-11623, 12521-12661, 13453-13664, 13824-13957, 15376-15478, 16855-16965, 17378-17495, 18535-18642, 21446-21541, 21999-22087, 23036-23247, 23546-23667, 24270-24461, 26287-26470, 26611-26747, 28068-28260, 32540-32709, 33112-33270, 34586-34828, 35156-35287, 36660-36763, 36934-37077, 37803-37921, 38017-38138, 40365-40493, 42618-42848, 43452-43578, 44836-44999, 48223-48269, and 49656-49779;

c) a contiguous span of at least 12 nucleotides of SEQ ID No 1 or the complements thereof, wherein said contiguous span comprises at least one BAP28-related biallelic marker selected from the group consisting of A1 to A58;

d) a contiguous span of at least 12 nucleotides of a nucleic acid sequence selected from the group consisting of SEQ ID Nos 2 and 3 or the complements thereof, wherein said contiguous span comprises at least 1 of nucleotide positions 1 to 4995 of SEQ ID No 2 or 3;

e) a contiguous span of at least 12 nucleotides of a nucleic acid sequence selected from the group consisting of SEQ ID Nos 2 and 3 or the complements thereof, wherein said contiguous span comprises at least 1 of nucleotide positions 1 to 2033, 2160 to 2348 and 2676 to 4995 of SEQ ID No 2 or 3;

f) a contiguous span of at least 12 nucleotides of a nucleic acid sequence selected from the group consisting of SEQ ID Nos 1-3 or the complements thereof, wherein said contiguous span comprises at least 1 of any one of the following ranges of nucleotide positions of:

(1) SEQ ID No 1: 1-2500, 2501-5000, 5001-7500, 7501-10000, 10001-12500, 12501-15000, 15001-17500, 17501-20000, 20001-22500, 22501-25000, 25001-27500, 27501-30000, 30001-32500, 32501-35000, 35001-37500, 37501-40000, 40001-42500, 42501-45000, 45001-47500, 47501-50000, 50001-50357, 50499-50963, 51257-52147, 52299-53234, 53394-53553,

53689-53837, 53943-54028, 54198-54740, 54896-55753, 55913-57385, 57495-58503, 58828-85946, 59355-85946, 86169-91228, and/or 91852 to 97662;

(2) SEQ ID No 2: 1 to 500, 501 to 1000, 1001 to 1500, 1501 to 2000, 2001 to 2500, 2501 to 3000, 3001 to 3500, 3501 to 4000, 4001 to 4500, 4501 to 4995, 5000 to 5500, 5501 to 6000, 6001 to 6500, and 6501 to 6782; and,

(3) SEQ ID No 3: 1 to 500, 501 to 1000, 1001 to 1500, 1501 to 2000, 2001 to 2500, 2501 to 3000, 3001 to 3500, 3501 to 4000, 4001 to 4500, 4501 to 4995, 5000 to 5500, 5501 to 6000, 6001 to 6500, 6501 to 7000, 7001 to 7500, 7501 to 7932; and

g) a nucleotide sequence selected from the group consisting of SEQ ID Nos 4, and 9-13; and,

h) a nucleotide sequence complementary to any one of the preceding nucleotide sequences; and

a polypeptide code comprising a contiguous span of at least 6 amino acids of SEQ ID No 5, wherein said contiguous span includes:

- at least 1 of the amino acid positions 1 to 1629 of the SEQ ID No 5; or,
- an amino acid selected from the group consisting of an asparagine at the amino acid position 1694 of SEQ ID No 5, a valine at the amino acid position 1854 of SEQ ID No 5, an asparagine at the amino acid position 1967 of SEQ ID No 5, a glutamic acid at the amino acid position 2017 of SEQ ID No 5, and an alanine at the amino acid position 2050 of SEQ ID No 5.

57. Use of a polynucleotide comprising a contiguous span of at least 12 nucleotides of the SEQ ID No 1 or the complementary sequence thereto for determining the identity of the nucleotide at a BAP28-related biallelic marker

58. Use according to claim 57 in a microsequencing assay, wherein the 3` end of said contiguous span is located at the 3` end of said polynucleotide and wherein the 3' end of said polynucleotide is located 1 nucleotide upstream of said BAP28-related biallelic marker in said sequence.

59. Use of according to claim 57 in a hybridization assay, wherein said span includes said PG1-related biallelic marker.

60. Use according to claim 57 in a specific amplification assay, wherein the 3' end of said contiguous span is located at the 3' end of said polynucleotide and said biallelic marker is present at the 3' end of said polynucleotide.

61. Use according to claim 57 in a sequencing assay, wherein the 3' end of said contiguous span is located at the 3' end of said polynucleotide.

62. Use according to claim 57, wherein said BAP28-related biallelic is a biallelic marker selected from the group consisting of A1 to A58, and the complements thereof.

63. Use according to claim 57, wherein said BAP28-related biallelic is a biallelic marker selected from the group consisting of A1 to A27, A34, A37 to A41, A43 to A49, A52, and A54 to A58, and the complements thereof.

64. Use according to claim 57, wherein said BAP28-related biallelic is a biallelic marker selected from the group consisting of A1, A4, 16, A30, A31, A42, A50, A51, and A53, and the complements thereof.